Functional mutants of the sequence-specific transcription factor p53 and implications for master genes of diversity

Michael A. Resnick* and Alberto Inga

Laboratory of Molecular Genetics, National Institute of Environmental Health Sciences, National Institutes of Health, Research Triangle Park, NC 27709

Communicated by Paul L. Modrich, Duke University Medical Center, Durham, NC, June 19, 2003 (received for review April 2, 2003)

There are many sources of genetic diversity, ranging from programmed mutagenesis in antibody genes to random mutagenesis during species evolution or development of cancer. We propose that mutations in DNA sequence-specific transcription factors that target response elements (REs) in many genes can also provide for rapid and broad phenotypic diversity, if the mutations lead to altered binding affinities at individual REs. To test this concept, we examined the *in vivo* transactivation capacity of wild-type human and murine p53 and 25 partial function mutants. The p53s were expressed in yeast from a rheostatable promoter, and the transactivation capacities toward >15 promoter REs upstream of a reporter gene were measured. Surprisingly, there was wide variation in transactivation by the mutant p53s toward the various REs. This is the first study to address directly the impact of mutations in a sequence-specific transcription factor on transactivation from a wide array of REs. We propose a master gene hypothesis for phenotypic diversity where the master gene is a single transcriptional activator (or repressor) that regulates many genes through different REs. Mutations of the master gene can lead to a variety of simultaneous changes in both the selection of targets and the extent of transcriptional modulation at the individual targets, resulting in a vast number of potential phenotypes that can be created with minimal mutational changes without altering existing protein–protein interactions.

transactivation | evolution | networks | mutation | promoter

The ability of organisms to diversify their phenotypes is important in many aspects of biology. Antibody variability is an example of a programmed cellular system for phenotypic diversity in higher eukaryotes that enables cells to generate a plethora of diverse antibodies in a short time (1). Nonprogrammed phenotypic diversification is found in many cancers where a variety of DNA and chromosomal changes can lead to cells that are able to commandeer the resources of the total organism (2). Accelerated diversification provides opportunities for organisms to adapt to acute challenging situations, to occupy new niches, and to lead to dramatically increased rates of evolution (see, for example, refs. 3 and 4).

Although the independent alteration of many genes is a generally accepted mode of phenotypic diversification, changes in a single gene could also provide for rapid diversification. For example, the inactivation of single genes that guard against mutation, such as those involved in mismatch repair (5), can lead to diversification via gross genome instability. The temporary enhancement of mutagenesis when mismatch repair is shut down provides for genetic/phenotypic diversification and the opportunity for bacteria to evolve and occupy new niches.

Another single-gene source of rapid phenotypic diversity, which we consider here, could be DNA sequence-specific transcription factors. Mutations in those transcription factors that target response elements (REs) in many genes could provide for rapid and broad phenotypic diversity, if the mutations led to altered binding affinities at individual REs. [We note that this diversity could occur without the proposed need (3) for gene duplications or allo-polyploidy.]

A formal way to express the extent of impact of a transcription factor gene "*X*" would be to identify all of the biological outcomes that can be affected by its product. For example, the corresponding DNA sequence-specific regulatory protein "p*X*" might interact with many REs, each one having a slightly different sequence. The ability of p*X* to induce transcription of a gene is determined in part by its ability to bind to a RE, referred to as "y." Binding can be strongly influenced by a RE sequence as well as other factors, such as posttranslational modifications of the protein, cofactors, and chromatin state at the RE site. The differences in binding will result in various amounts of transcription of the individual genes ($Y_1, Y_2, \ldots, Y_c$) directly regulated by the protein p*X* through interaction with the respective REs ($y_1, y_2, \ldots, y_c$). The set of transcriptional outcomes can be denoted as $T_{X,yc}$.

Mutational changes in the regulatory protein p*X* could result in new sets of expression outcomes and, therefore, phenotypic diversity. As an example, $T_{\Delta x,yc}$ would be the set of outcomes with a deletion mutation Δx where the transcription of the corresponding genes becomes independent of p*X*. [The differences in expressions levels for the individual target genes could be detected with expression microarrays (6).] If the protein could tolerate various in-frame mutational changes x1, x2, etc., that alter its activity, then new sets of outcomes, or phenotypic states, $T_{x1,yc}$, $T_{x2,yc}$, etc., would be possible depending on the nature of the individual mutations. The consequences of mutations in gene *X* would not necessarily lead to the same transactivation effects from the various REs. There might be increases, decreases, or no change, and the levels of change may be specific to the individual REs so that new patterns of responses are developed. The extent of phenotypic diversity would be determined by the number of mutations in gene *X* that can lead to new sets of phenotypes. In addition to mutations, there could be transient variations caused by changes in either the levels and/or activity of p*X*, referred to as "Xe," that might lead to additional sets of expression outcomes ($T_{Xe,yc}$) and, therefore, greatly increase the number of phenotypes. Thus, a combination of functional mutants of a DNA sequence-specific transcription factor and variations in expression could result in a vast number of phenotypic outcomes. A single gene could then be considered a master gene for phenotypic diversity. If such master genes exist, they would provide for the economic creation of multiple phenotypes.

The human p53 gene provides a good system for investigating the concept of master genes for genetically derived phenotypic diversity. The p53 gene product is a DNA sequence-specific transcription factor which, as a homotetramer, controls the expression of >50 downstream genes through direct binding with REs (7). (Note that, depending on the context, p53 may refer to the gene *p53* or the protein p53.) The sequence of each RE is

Abbreviations: RE, response element; DNA-BD, DNA-binding domain.

*To whom correspondence should be addressed. E-mail: resnick@niehs.nih.gov.

www.pnas.org/cgi/doi/10.1073/pnas.1633803100

related to a degenerate 20-bp consensus sequence, and deviations from the consensus are common (8) (see Table 1, which is published as supporting information on the PNAS web site, www.pnas.org). The RE sequence can affect the level and/or the kinetics of transactivation. There are likely to be hundreds of genes and, therefore, specific REs that are direct p53 targets (9, 10). In general, the target genes of p53 can be grouped into categories of biological activities that include apoptosis, growth arrest, DNA repair, and checkpoint responses. There are also direct interactions with many proteins (including DNA repair/recombination gene products such as BRCA1, RAD51, XPB, XPD, etc.) that may provide further opportunities for diverse responses (11). The wild-type p53 contributions to cellular phenotypes can be considered as the sum of all sets of target gene functions and interactions with other proteins. Levels and timing of expression of the p53 target genes can contribute to the differences in phenotypes (10, 12, 13). The transcriptional outcomes, denoted as $T_{p53,yc}$, correspond to the binding and transactivation of p53 at the response elements (y_1, y_2, . . . , y_c).

Because of variations in expression and stresses, the "on" (wild type) allele of p53 can lead to many different phenotypes, whereas the "off" (mutant) allele would lead to a much smaller number. However, there are p53 mutants (i.e., p53-1, p53-2, etc.) that retain function and alter binding to REs (14, 15). The fact that nearly 80% of tumor-associated p53 mutations are missense and that over a thousand different single amino acid changes have been identified (16) suggests that the category of mutations with partial or altered function could be large. In principle, a p53 allele could cause various changes in transcription from REs y_1, y_2, . . . , y_c, of downstream genes, such as the p21- and bax-binding sites, resulting in new sets of transcriptional outcomes $T_{p53-1,yc}$, $T_{p53-2,yc}$, etc. Changes in the pattern of target gene activation, as well as levels of transcription, would in essence correspond to changes in phenotypes. By examining the consequences of amino acid changes in the sequence-specific DNA-binding region of p53 on transactivation from many target REs using an isogenic yeast-based system, we establish that a single gene, such as p53, could indeed be a source of vast variation in phenotypes, i.e., a master gene of diversity.

Materials and Methods

Construction of Isogenic Yeast Strains Containing p53-Responsive Promoters Upstream of Reporter Genes. The *delitto perfetto in vivo* site-specific mutagenesis system (17) was used to generate isogenic yeast strains containing different p53 REs at the natural *ADE2* locus. For the *ADE2*-based strains, we used the previously described *yAFM-CORE* isolate (18). A desired p53 RE was introduced, replacing the CORE cassette (CO: counterselectable, *KLURA3*; RE: reporter, *KanMX4*, G418 resistance gene) by using appropriate oligonucleotides. The luciferase cDNA was inserted in place of the *ADE2* ORF by using a two-step recombination procedure (17). This strain, *yLFM*, was prepared for insertion of p53 REs by inserting a CORE cassette within the minimal promoter as for yAFM. Thus, *yAFM-REs* and *yLFM-REs* differ only in the gene reporter ORFs and the RE sequences. All CORE replacements were confirmed by PCR and DNA sequencing. The sequence of primers and additional details on strain construction are available on request.

Phenotypic and Quantitative Transactivation Assays at Variable Levels of p53 Expression. The transactivation of the *ADE2* gene results in a colony color change from red to pink or white (18). Plasmid pLS89 (obtained from Richard Iggo, Oncogene Group, Institut Suisse de Recherches Experimentales sur le Cancer, Epalinges, Switzerland) is a centromeric expression vector that contains the various p53 cDNAs under the inducible *GAL1-10* promoter. Plasmids expressing the BRCA-1 associated p53 mutations were obtained from Paola Monti and Gilberto Fronza [Laboratory of Mutagenesis, Istituto Scientifica (per lo studio e la cura dei) Tumori, Genoa, Italy]. The vectors were introduced into the yeast strains by LiAc transformation (19) and selected on glucose plates lacking tryptophan and containing high adenine to overcome the growth defect caused by incomplete adenine biosynthesis. The transactivation assay was then performed on plates with low adenine (5 mg/liter) to allow color detection by using three independent transformants of each *yAFM*-RE strain. Expression of p53 was modulated by using a combination of different sugars in the plate assay (18). Plates were incubated for 3 days at 30°C before scoring colony color. p53 expression was examined by Western blot using pAb1801 and DO-1 monoclonal antibodies (Santa Cruz Biotechnology) (20).

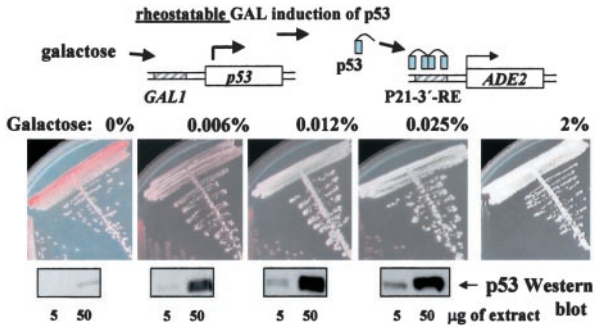

Fig. 1. System for examining the transactivation capacity of wild-type and mutant p53s toward various response elements (REs). The *p53* cDNA is expressed in yeast under the control of the *GAL1,10* promoter, whose induction is regulated by varying the amounts of galactose. The *ADE2* gene is controlled by p53 because a p53 RE (P21–3′ in the example) was inserted within a modified minimal promoter. Transactivation is assessed by colony color and depends on the amount of galactose in the synthetic medium (cells grown at five different concentrations are shown); the synthetic medium also contains 2% raffinose, which results in a basal level of p53 expression (18). A Western blot showing the amount of p53 protein is included.

To quantitate p53-dependent *ADE2* transactivation, the ratio between *ADE2* and *ACT1* (actin) mRNAs was determined by using quantitative RT-PCR approaches starting from total RNA obtained from liquid cultures after 24 h of growth. RNA isolation, reverse transcription, and quantification by PCR in real time were performed as described (18). To quantify the luciferase activity, the p53 transformants were recovered after 24 h of growth, and lysed by using lysis reagent (Promega) and glass beads (0.5 mm, acid washed, Sigma) by cycles of vortexing and incubation on ice. Soluble proteins were purified by centrifugation and quantified by using the Bradford method (Bio-Rad). Luciferase activity was measured with a multilabel plate counter Wallac Victor (Perkin–Elmer) by using luciferase assay reagent (Promega).

Results

Assessing Altered Spectra and Levels of Transcriptional Responses in p53 Mutants. Recently, we developed a system in the yeast *Saccharomyces cerevisiae*, described in Fig. 1, that assesses the transactivation capacity of human p53 toward >30 REs derived from human genes, where the REs are placed upstream from a reporter gene such as the color marker *ADE2* (18). We used this system, which is based on the capability for highly regulatable (i.e., "rheostatable") expression of p53, to demonstrate a 1,000-fold range in transactivation capacities between the REs by wild-type p53 (see Fig. 5, which is published as supporting information on the PNAS web site). The opportunity to vary the p53 expression levels enabled us to identify mutations that had altered abilities to transactivate from some of these REs (20, 21).

The initial observations with a few functionally altered p53 mutations and a small number of REs led us to investigate the

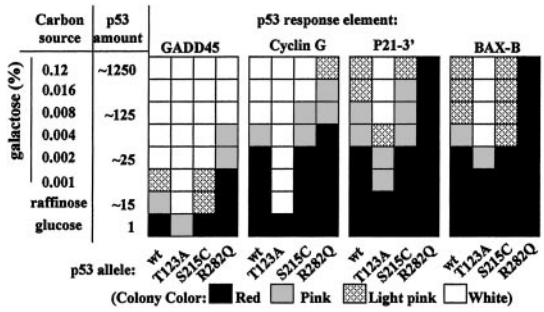

Fig. 2. Examples of p53-induced transactivation by using the ADE2 reporter gene assay and different levels of p53 expression. Wild-type human p53 and three functional mutants were examined at different levels of induction for their ability to induce transcription at four REs, as determined by change in colony color. The relative induction of p53 protein is shown. The supertrans T123A mutant showed enhanced activity with all four REs, whereas S215C, a mutant associated with familial breast cancer, exhibited subtle defects and the tumor hotspot mutant R282Q had no activity with the weak P21–3′ and BAX-B REs.

ability of 25 mutations (i.e., p53-1, p53-2, etc.) in the DNA-binding domain (DNA-BD) of p53 to transactivate from over 15 REs (i.e., $y_1, y_2, \ldots, y_c$) by using the above system. The sequences of the REs, which differ by 1–4 bases from the consensus sequence (Table 1), were taken from various human p53 targets that include apoptosis, DNA repair, and cell cycle response genes. With this reporter system, it is possible to identify new sets of outcomes, corresponding to interactions of mutant p53 proteins with various target sequences at different levels of p53 expression. This is particularly important because the level of p53 in human cells is highly responsive to stress signals. Along this line, expression of subtle p53 mutants may result in outcomes that are comparable to wild type at high, but not low, expression levels, as appears to be the case for some p53 mutations associated with familial breast cancer (14, 18).

Many p53 Mutants Alter the Spectrum and Level of Transactivation Responses. Differences in mutant p53 transactivation capacity were assessed by determining the ability of mutant and wild-type p53s to produce a change in colony color from red (no induction from the RE), to light pink and pink (mild induction), to white (strong induction) when different levels of the galactose inducer were added to the raffinose-containing medium (ranging from 0% to 2% galactose, resulting in basal to maximal p53 expression) (18, 22). Although the mutants are functional, they differ considerably in their ability to induce transcription from the various REs. An example of this analysis is presented in Fig. 2. To verify responses with the color reporter system, we replaced the *ADE2* ORF with a luciferase (LUC) reporter gene in the same chromosomal locus. Overall, there is a strong correlation between visual and quantitative reporter assays (see Fig. 6, which is published as supporting information on the PNAS web site).

The relative transactivation capacities of the 25 mutant p53s with respect to wild-type p53 toward 15 REs are summarized in Fig. 3*A*, where they are presented in a format similar to that found for expression microarrays. Results with additional REs for a subset of the mutants are presented in Fig. 3*B*. Increased activity relative to wild-type p53 is depicted by shades of color that extend from yellow (similar to wild-type response) to red (≈16 times higher activity). Increases are inversely related to the amount of galactose inducer (i.e., level of p53) required to turn on the reporter with mutant as compared with wild-type p53. Conversely, if more mutant p53 protein relative to wild type is required for induction from a particular RE, the reduced transactivation capacity is depicted in different shades of green, with weaker p53s being colored in darker green. Loss-of-function is depicted in blue.

Among the 25 mutants examined, half were located in the L1 loop domain. Mutants 121F through 124Y and 288K were identified in two separate screens for p53 alleles that could lead to enhanced transactivation (supertrans) at the RGC or p21–5′ RE (ref. 21 and unpublished results). Mutant 288D was included as a comparison with 288K. In the three-dimensional structure, the 121–124 amino acids of the L1 loop face the H2 helix, which is also involved in DNA contacts (amino acids 277, 280, and 283) (23) and may contribute to L1 loop/H2 helix positioning. Position 120 is a DNA contact site at the tip of the loop, but, unlike all other DNA contact amino acids, it is not mutated in tumors (24). Amino acids 117 and 119 are at the opposite side of the L1 loop (23). Mutants 122A, 125K and R, 277R and W, and 279R are toxic at high expression levels (ref. 20 and unpublished results). Mutant 279E has a different side chain from the toxic allele 279R. Mutants 150I, 199R, 202S, 215C and 283H were found in familial (BRCA1/2-associated) breast cancer patients (14) and showed wild type levels of transactivation or subtle defects (18, 25). Mutant 282Q retains activity toward p21–5′ but not the bax RE at high expression (25). Also presented are results with wild-type murine p53, whose activity was not distinguishable from human p53 (Fig. 3*A*), and the mutant T122L, (corresponding to human 125L) which is a UV-induced hotspot mutation in skin cancers from mice deficient in global genome repair (XPC −/− knockout) (26, 27).

Clearly, single amino acid changes in p53 can lead to considerable diversity in the spectrum of responses from the different REs, rather than simple general increases or decreases in transcription. The same p53 mutant can result in increases, decreases, and comparable levels of transactivation relative to wild type. Also, there do not appear to be specific regional effects. For example, although the 120E DNA contact mutation knocked out transcription for all of the REs, a change in amino acid at 119 had a subtle effect and a mutation at 121 enhanced transactivation for all REs. Based on the spectra and levels of transactivation, all mutants appear to be functionally different. Four broad classes of responses can be identified: (*i*) decrease/loss-of-function, found with the tumor mutants G279E, R282Q, N288D, and the contact-site mutant K120E; (*ii*) subtle changes, observed with mutants reported in familial breast cancer; (*iii*) altered specificity, with both severe increase, decrease, and loss of function toward specific REs, found among toxic mutants and the murine T122L hotspot; and (*iv*) supertransactivation. Although each p53 allele in the latter class differed in its impact on the individual REs, examples of supertransactivation could be found for each RE. Based on Western blot analysis (ref. 21 and unpublished results) the differences in mutant p53 activities cannot be ascribed to intrinsic differences in stability. Along this line, the present results that demonstrate dramatic shifts in the spectra and strength of transactivation at many REs with single amino acid changes in p53 suggest that simple structural and energy models for predicting binding efficiency of regulatory proteins do not explain the strength of binding to REs (9, 18).

Discussion

In this study, we demonstrate that single amino acid changes can have diverse effects on the ability of p53 to induce transcription from many REs. These effects were revealed by varying the expression of p53. Following the description presented in the Introduction, the spectrum of responses for each mutant p53-1, p53-2, etc., at various levels of expression can be considered as different sets of transcriptional outcomes where mutants p53-1, p53-2. . . , etc., have outcomes $T_{(p53\text{-}1)e1,yc}$; $T_{(p53\text{-}2)e2,yc}$, . . . , $T_{(p53\text{–}2)e1,yc}$; $T_{(p53\text{–}2)e2,yc}$, . . . , etc.

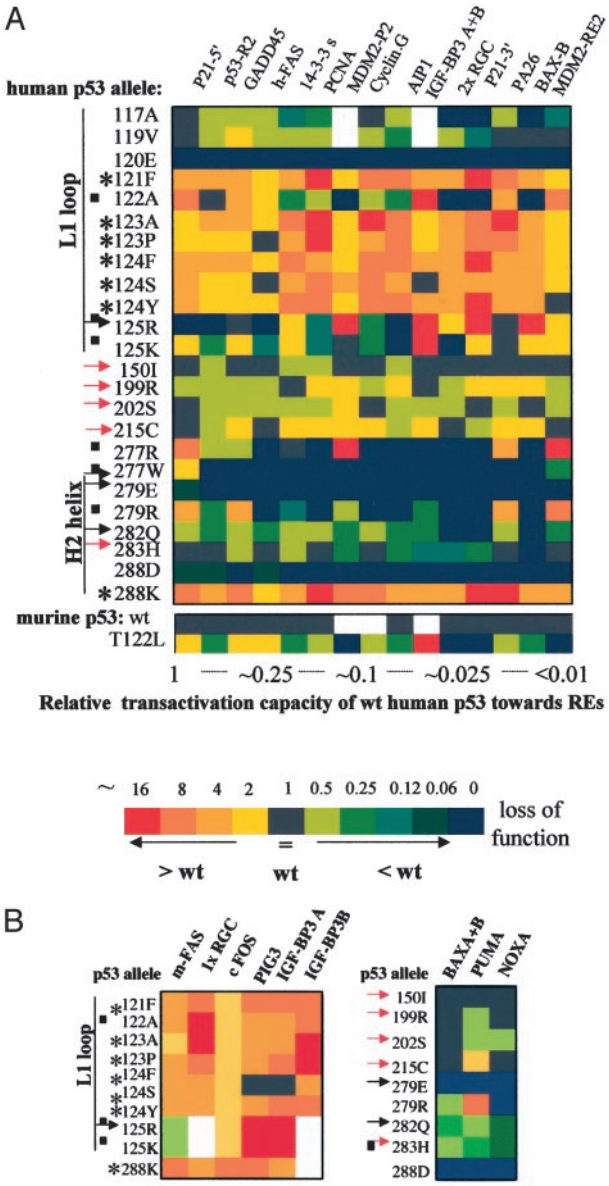

Fig. 3. Relative transactivation capacity of p53 alleles with mutations in the DNA-BD. (*A*) Twenty-four human p53 mutations were examined in 15 isogenic yeast strains, each containing a different p53 RE that regulates expression of the *ADE2* reporter gene. The REs are ranked from left to right according to their decreasing transactivation capacity with wild-type p53 (indicated at the bottom; also see Fig. 5). The p53 mutants are ordered according to their position in the primary sequence. The topological domains are shown. Red and black arrows identify mutations associated with familial breast cancer and sporadic cancers, respectively. The transactivation capacity of each allele toward each RE was determined by using variable expression of p53 under the *GAL1* promoter and compared with the activity of wild-type p53. The relative transactivation capacities of mutants with respect to wild type is presented in a form similar to that for expression microarrays, with red indicating greatly increased, green indicating greatly decreased, blue indicating loss of function, and black equal to wild type. The quantification is based on the amounts of p53 protein required for transactivation with wild type or with a mutant allele. This was derived from the minimal amount of galactose required for transactivation to occur. p53 mutations with enhanced transactivation toward most REs are classified as supertrans (asterisks). Alleles with greatly altered patterns of specificity, including increase, decrease, and loss-of-function, are designated with squares. Alleles associated with sporadic and familial breast cancer are designated with black and red arrows, respectively. The transactivation capacity of murine wild-type p53 and of the T122L mutant (corresponding to T125L in the human protein) is also shown. The human and mouse wild-type p53s were not distinguishable in this assay. (*B*) In addition to the 15 REs examined above, seven supertrans and three altered specificity p53 mutants were characterized with six additional REs (*Left*). Although four REs (1× RGC through IGFBP3-A) are weak with wild-type p53 and one (IGFBP3-B) is not functional (see Fig. 5), enhanced transactivation was observed with most elements. Interestingly, the altered specificity mutants 125R and 125K showed reduced capacity with the strong RE of *m*-FAS but greatly enhanced activity with the weak PIG3 and IGFBP3-A REs. Three additional apoptotic REs were also investigated with nine more p53 alleles associated with cancer (*Right*). Mutant G279R showed enhanced activity with PUMA RE, reduced activity with BAX A+B, and lack of function with NOXA. The BAX A+B contains two adjacent p53 REs (see Table 1).

Clearly, a small number of mutations in a DNA sequence-specific transcription factor could provide many new sets of outcomes, or potentially different phenotypes, especially when expression levels of the factor are considered. Given the large number of mutations in p53 that retain function (this study and unpublished results) and various levels of p53 expression that occur normally or in response to stress, there could be hundreds of unique phenotypic states in terms of overall cellular transcription. Further complexity in phenotypic diversity would result if changes such as posttranslational and RNA processing modifications are included. The impact of the individual mutations cut across target sequences in biologically related groups of genes (Table 1), suggesting that they could result in a variety of biological consequences.

These observations are relevant to the etiology of cancer, which can be considered a microevolution disease. Alterations in the tumor suppressor p53 are intrinsic to most cancers where modifications include changes in expression or stability, loss, or altered functions such as those described here (11). It is interesting to consider that various p53 functional mutations may relate to specific types of cancers (28, 29). For example, in UV-induced skin cancers of mice that are homozygous for a global DNA excision repair defect and hemizygous for p53, the T122L mutation is frequently induced (26). We have suggested that the marked perturbations in transcriptional responses shown in Fig. 3 may be selected in the particular mutant background (27). For the case of familial breast cancer-associated tumors, where there is a nearly 75% association of p53 mutations, there is a paucity of the typical p53 mutant hotspots and several p53s have only subtle defects (14, 18, 30).

Among the nearly 17,000 clinically identified p53 mutations associated with cancer, ≈1,200 different single amino acid changes have been reported in the IARC database (24). It is not known how many missense or small in-frame deletion p53 mutations retain function, although data from functional assays or structural modeling suggest that their number could be high (between 20% and 50% of amino acid changes; refs. 16 and 25 and our unpublished results). It is likely that many of these functional mutants alter the p53 transcriptional responses, resulting in a variety of phenotypic outcomes. The approach that we have taken may prove useful in better evaluating the prognostic value of p53 status in tumors, including the responsiveness to therapy (31), in which case it would be important to include functionality data in any mutant p53 database.

Mouse and Human p53s and RE Targets. The comparison between human and murine wild-type p53s, which are 89% identical in the DNA-BD, did not reveal differences in transactivation capacity in the yeast assay (Fig. 3). This is consistent with the high degree of similarity between the three-dimensional structures of the DNA-BDs revealed by crystal structure analysis (32) and the ability to exchange the mouse and human DNA-BDs in Hupki knock-in mice (33). Differences in transactivation capacity might be detectable with REs that are highly diverged from the consensus.

The results with the mouse and human p53s may indicate that the development of the bulk of the p53 response pathways preceded the evolutionary branching that led to the origin of rodents and primates. Differences in the repertoire of p53-modulated transcriptional patterns have evolved, in part, through changes in promoter sequences of the individual target genes. For example, in the case of the *DDB2* gene, which is responsive to p53 in human cells but not mouse cells, a p53 RE was identified only in the human promoter, although the mouse and human promoters share high overall sequence identity (34). Given our findings of wide variations in transactivation capacity among p53 REs (this study and ref. 18), it would be interesting to systematically compare the actual sequence of the REs within the regulatory regions of p53 target genes in different species and their ability to support transactivation by the human and mouse p53s. It would also be worthwhile to compare the transactivation capacities between evolutionary distant p53 homologues. Interestingly, the p53-related genes p63 and p73, which are ≈65% identical to p53 in the DNA-BD (35), exhibit differences in transactivation capacity in our yeast-based assay (unpublished results).

Mutations in DNA Sequence-Specific Transcription Factors and Evolution of Regulatory Networks. Here we attempt to address the role of specific amino acid changes in a DNA sequence-specific transcription factor in terms of the spectra and levels of transcriptional responses from the REs of many genes in a eukaryotic cell environment under completely isogenic conditions. By using the p53 gene as an example, we have established that changes in a single gene could potentially be a powerful source of genetically determined phenotypic diversity. Diversity is accomplished through mutations in a transcription factor resulting in differential changes in the levels of expression of many target genes. This mechanism for generating diversity via a single gene contrasts sharply with more traditional models, where diversity and evolution proceed through the acquisition of mutations in many genes, either at coding or regulatory sequences.

It is interesting that proteins that regulate the expression and/or activity of several genes may account for ≈5% of coding sequences (36). Changes in these proteins could provide for evolution of regulatory networks (37, 38). Transcriptional networks provide for the appropriate and coordinated expression of groups of genes. Loss of genes such as Forkhead/winged-helix genes, Pax genes, prox1, p63, SMADs, APC, IkB, or NF-κB result in dramatic changes in transcriptional patterns leading to developmental, proliferation, and differentiation defects (39–44). Similarly, loss of the tumor suppressor p53 results in specific modifications in the response of many growth-restraint and apoptosis genes (45). Although loss of a single gene product in a regulatory network does not provide for diversity, because one mutation leads to a specific set of changes, it would be interesting to examine the consequences of several mutations in various DNA sequence-specific transcription factors, such as NF-κB, on the transcription of the respective target REs by using the approach used in this study with p53. There are few examples where the consequences of amino acid changes in a DNA sequence-specific transcription factor have been analyzed in terms of response of individual downstream targets. Most notable is the study of the AREA gene of *Aspergillus nidulans*, where specific amino acid changes altered pathways of nitrogen metabolism, although levels of expression of the AREA gene were not examined (46).

Repressors could also be master genes of diversity. For example, the lexA gene product can act at >20 sequence-related sites to repress genes associated with various DNA damage and nutritional stresses. The mutation lexA41 results in a change in the response associated with only a subset of repressible sites (47).

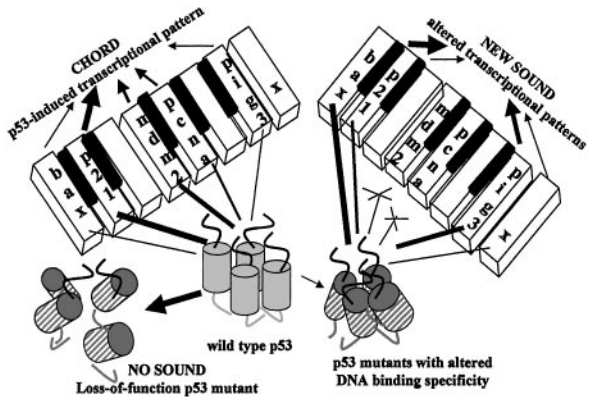

Fig. 4. The piano analogy for the master gene of diversity hypothesis based on p53 responses. A master gene such as p53 controls the expression of many genes, similar to hands playing notes on a piano. The p(53) piano has many keys corresponding to direct target genes (>50), and it is played by the ''fingers'' of the p53 ''hand'' via interaction with promoter REs. This creates a ''chord'' where each note (key) may have a different intensity (i.e., level of transactivation). In the case of p53 mutants that are completely nonfunctional, none of the target genes are activated and ''no sound'' will emanate. However, for p53 mutants retaining partial function or exhibiting altered DNA-binding specificity, the p53 hand can have some inactive ''fingers'' that do not strike keys, and some fingers that are stronger (bold type) or weaker (smaller and lighter type). It is even possible that additional new keys (denoted by ''x'') will be struck as would be the case for recognition of related REs. These changes can result in new chords and sounds with novel biological consequences.

Intrinsic to the concept of master genes of diversity is the opportunity to evolve without the constraint of coevolution of interacting proteins (48). Most cellular activities are accomplished by groups of complexes of proteins, and changes in one of the proteins within a complex may require changes in other proteins to maintain function of the complex. In the master gene hypothesis, there is no need for coevolution because there does not need to be a dramatic change in the ability of proteins to interact. Thus, there may be greater opportunities for accelerated evolution.

The Master Gene Hypothesis and Piano Analogy. The concept of single master genes of diversity leads us to describe a model that uses, by way of analogy, chords produced from a musical instrument such as a piano. In this model, which is presented in Fig. 4, each ''key'' of the piano corresponds to a gene whose expression is determined by the interaction of its REs with a DNA sequence-specific transcription factor or repressor. The transcription factor or repressor is considered the ''hand'' that plays the notes. As a consequence of the fingers of the hand touching the keys, a chord is produced that represents the biological response. The fingers of the hand may exert different pressures, and this leads to various intensities of the individual notes (for example, see Fig. 5). Changes in the hand, corresponding to specific mutations, can lead to a variety of changes in the sound emanating from each key. A null mutation corresponds to a nonfunctional hand. New chords and sounds would result if there are changes in the fingers capable of playing notes or if there are differences in the strength of the fingers. Thus, with many notes and intensities, there are a vast number of possible chords and sounds of chords. In this analogy, the new chords correspond to potentially different phenotypes. The piano interpretation of the master gene hypothesis can be applied to many transcription factors that have multiply diverged target sequences, such as NF-κB, PHA-4, pax-6, prox1, and the AREA gene.

Transcriptional networks could also evolve through alterations in

the promoter regions of genes (49), so as to put them under the control of transcription factors (e.g., the *DDB2* gene described above), or change the robustness of transcriptional control by a given factor (38, 50, 51). In the piano analogy, this would correspond to the addition or loss of keys. This might be the case for polymorphisms in target sequences (52, 53) or for neutral mutations that can lead to the relatively rapid establishment of new promoter response elements (49). The combination of mutations in promoter regions, mutation in regulatory genes, and selection would be expected to provide opportunities for coordinated rapid evolution of multigene systems. This could be important in the "acquisition of genes for organ development" as described for the development of the pharynx in *Caenorhabditis elegans* (50) or transcriptional network evolution, where genes that are subject to transcriptional controls are changed (37, 38).

Note Added in Proof. These results are supported by the recent report of Kato *et al*. (54). The authors have examined >2,000 p53 mutants in yeast for ability to transactivate under conditions of high expression from eight different tandem response element arrays and found many mutants that were "on" for some response elements and "off" for others. The system developed in our work has enabled us to address a wide variety of effects of mutations such as subtle changes, supertransactivation, and even toxicity.

We thank Thomas Darden, Francesca Storici, Dmitry Gordenin, Jack Taylor, and Jan Drake at National Institute of Environmental Health Sciences and Dr. Gilberto Fronza at Istituto Scientifico (per lo studio e la cura dei) Tumori (Genoa, Italy) for materials, advice, and comments on the manuscript. This work was partially supported by Department of Defense Grant DAMD 17-01-0579 (to M.A.R.). A.I. is supported by a National Institute of Environmental Health Sciences Intramural Research Award.

1. Gearhart, P. J. (2002) *Nature* **419,** 29–31.
2. Hanahan, D. & Weinberg, R. A. (2000) *Cell* **100,** 57–70.
3. Barrier, M., Robichaux, R. H. & Purugganan, M. D. (2001) *Proc. Natl. Acad. Sci. USA* **98,** 10208–10213.
4. Lucht, J. M., Mauch-Mani, B., Steiner, H. Y., Metraux, J. P., Ryals, J. & Hohn, B. (2002) *Nat. Genet.* **30,** 311–314.
5. Giraud, A., Radman, M., Matic, I. & Taddei, F. (2001) *Curr. Opin. Microbiol.* **4,** 582–585.
6. Yoon, H., Liyanarachchi, S., Wright, F. A., Davuluri, R., Lockman, J. C., de la Chapelle, A. & Pellegata, N. S. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 15632–15637.
7. Levine, A. J. (1997) *Cell* **88,** 323–331.
8. el-Deiry, W. S. (1998) *Semin. Cancer Biol.* **8,** 345–357.
9. Hoh, J., Jin, S., Parrado, T., Edington, J., Levine, A. J. & Ott, J. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 8467–8472.
10. Zhao, R., Gish, K., Murphy, M., Yin, Y., Notterman, D., Hoffman, W. H., Tom, E., Mack, D. H. & Levine, A. J. (2000) *Genes Dev.* **14,** 981–993.
11. Prives, C. & Hall, P. A. (1999) *J. Pathol.* **187,** 112–126.
12. Szak, S. T., Mays, D. & Pietenpol, J. A. (2001) *Mol. Cell. Biol.* **21,** 3375–3386.
13. Yu, J., Wang, Z., Kinzler, K. W., Vogelstein, B. & Zhang, L. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 1931–1936.
14. Smith, P. D., Crossland, S., Parker, G., Osin, P., Brooks, L., Waller, J., Philp, E., Crompton, M. R., Gusterson, B. A., Allday, M. J. & Crook, T. (1999) *Oncogene* **18,** 2451–2459.
15. Ludwig, R. L., Bates, S. & Vousden, K. H. (1996) *Mol. Cell. Biol.* **16,** 4952–4960.
16. Martin, A. C., Facchiano, A. M., Cuff, A. L., Hernandez-Boussard, T., Olivier, M., Hainaut, P. & Thornton, J. M. (2002) *Hum. Mutat.* **19,** 149–164.
17. Storici, F., Lewis, L. K. & Resnick, M. A. (2001) *Nat. Biotechnol.* **19,** 773–776.
18. Inga, A., Storici, F., Darden, T. A. & Resnick, M. A. (2002) *Mol. Cell. Biol.* **22,** 8612–8625.
19. Gietz, R. D., Schiestl, R. H., Willems, A. R. & Woods, R. A. (1995) *Yeast* **11,** 355–360.
20. Inga, A. & Resnick, M. A. (2001) *Oncogene* **20,** 3409–3419.
21. Inga, A., Monti, P., Fronza, G., Darden, T. & Resnick, M. A. (2001) *Oncogene* **20,** 501–513.
22. Flaman, J. M., Frebourg, T., Moreau, V., Charbonnier, F., Martin, C., Chappuis, P., Sappino, A. P., Limacher, I. M., Bron, L., Benhattar, J., *et al.* (1995) *Proc. Natl. Acad. Sci. USA* **92,** 3963–3967.
23. Cho, Y., Gorina, S., Jeffrey, P. D. & Pavletich, N. P. (1994) *Science* **265,** 346–355.
24. Olivier, M., Eeles, R., Hollstein, M., Khan, M. A., Harris, C. C. & Hainaut, P. (2002) *Hum. Mutat.* **19,** 607–614.
25. Campomenosi, P., Monti, P., Aprile, A., Abbondandolo, A., Frebourg, T., Gold, B., Crook, T., Inga, A., Resnick, M. A., Iggo, R. & Fronza, G. (2001) *Oncogene* **20,** 3573–3579.
26. Reis, A. M., Cheo, D. L., Meira, L. B., Greenblatt, M. S., Bond, J. P., Nahari, D. & Friedberg, E. C. (2000) *Cancer Res.* **60,** 1571–1579.
27. Inga, A., Nahari, D., Velasco-Miguel, S., Friedberg, E. C. & Resnick, M. A. (2002) *Oncogene* **21,** 5704–5715.
28. Ribeiro, R. C., Sandrini, F., Figueiredo, B., Zambetti, G. P., Michalkiewicz, E., Lafferty, A. R., DeLacerda, L., Rabin, M., Cadwell, C., Sampaio, G., *et al.* (2001) *Proc. Natl. Acad. Sci. USA* **98,** 9330–9335.
29. Miyaki, M., Iijima, T., Ohue, M., Kita, Y., Hishima, T., Kuroki, T., Iwama, T. & Mori, T. (2003) *Gut* **52,** 304–306.
30. Greenblatt, M. S., Chappuis, P. O., Bond, J. P., Hamel, N. & Foulkes, W. D. (2001) *Cancer Res.* **61,** 4092–4097.
31. Soussi, T. & Beroud, C. (2001) *Nat. Rev. Cancer* **1,** 233–240.
32. Zhao, K., Chai, X., Johnston, K., Clements, A. & Marmorstein, R. (2001) *J. Biol. Chem.* **276,** 12120–12127.
33. Luo, J. L., Yang, Q., Tong, W. M., Hergenhahn, M., Wang, Z. Q. & Hollstein, M. (2001) *Oncogene* **20,** 320–328.
34. Tan, T. & Chu, G. (2002) *Mol. Cell. Biol.* **22,** 3247–3254.
35. Yang, A., Kaghad, M., Caput, D. & McKeon, F. (2002) *Trends Genet.* **18,** 90–95.
36. Tupler, R., Perini, G. & Green, M. R. (2001) *Nature* **409,** 832–833.
37. Rifkin, S. A., Kim, J. & White, K. P. (2003) *Nat. Genet.* **33,** 138–144.
38. Sengupta, A. M., Djordjevic, M. & Shraiman, B. I. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 2072–2077.
39. Mears, A. J., Jordan, T., Mirzayans, F., Dubois, S., Kume, T., Parlee, M., Ritch, R., Koop, B., Kuo, W. L., Collins, C., *et al.* (1998) *Am. J. Hum. Genet.* **63,** 1316–1328.
40. Chao, L. Y., Mishra, R., Strong, L. C. & Saunders, G. F. (2003) *Hum. Mutat.* **21,** 138–145.
41. Hong, Y. K., Harvey, N., Noh, Y. H., Schacht, V., Hirakawa, S., Detmar, M. & Oliver, G. (2002) *Dev. Dyn.* **225,** 351–357.
42. Celli, J., Duijf, P., Hamel, B. C., Bamshad, M., Kramer, B., Smits, A. P., Newbury-Ecob, R., Hennekam, R. C., Van Buggenhout, G., van Haeringen, A., *et al.* (1999) *Cell* **99,** 143–153.
43. Attisano, L. & Wrana, J. L. (2002) *Science* **296,** 1646–1647.
44. Darnell, J. E., Jr. (2002) *Nat. Rev. Cancer* **2,** 740–749.
45. Vogelstein, B., Lane, D. & Levine, A. J. (2000) *Nature* **408,** 307–310.
46. Wilson, R. A. & Arst, H. N., Jr. (1998) *Microbiol. Mol. Biol. Rev.* **62,** 586–596.
47. Peterson, K. R. & Mount, D. W. (1987) *J. Mol. Biol.* **193,** 27–40.
48. Fraser, H. B., Hirsh, A. E., Steinmetz, L. M., Scharfe, C. & Feldman, M. W. (2002) *Science* **296,** 750–752.
49. Stone, J. R. & Wray, G. A. (2001) *Mol. Biol. Evol.* **18,** 1764–1770.
50. Gaudet, J. & Mango, S. E. (2002) *Science* **295,** 821–825.
51. Zhou, Y., Zheng, J. B., Gu, X., Li, W. & Saunders, G. F. (2000) *Gene* **245,** 319–328.
52. Goldfeld, A. E., Leung, J. Y., Sawyer, S. A. & Hartl, D. L. (2000) *Gene* **261,** 19–25.
53. Contente, A., Dittmer, A., Koch, M. C., Roth, J. & Dobbelstein, M. (2002) *Nat. Genet.* **30,** 315–320.
54. Kato, S., Han, S.-Y., Liu, W., Otsuka, K., Shibata, H., Kanamaru, R. & Ishioka, C. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 8424–8429.